

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

#3 PCT
PHO 9956/1/01
8/7

Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

REC'D 01 SEP 2000	
WIPO	PCT

Bescheinigung

Certificate

Attestation

EJ4

EP 00 / 6436

Die angehefteten Unterla-
gen stimmen mit der
ursprünglich eingereichten
Fassung der auf dem näch-
sten Blatt bezeichneten
europäischen Patentanmel-
dung überein.

The attached documents
are exact copies of the
European patent application
described on the following
page, as originally filed.

Les documents fixés à
cette attestation sont
conformes à la version
initialement déposée de
la demande de brevet
européen spécifiée à la
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

99890234.0

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

Hatten

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 07/03/00
LA HAYE, LE

THIS PAGE BLANK (USPTO)



Europäisches
Patentamt

European
Patent Office

Office européen
des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.:
Demande n°: 99890234.0

Anmeldetag:
Date of filing:
Date de dépôt: 08/07/99

Anmelder:
Applicant(s):
Demandeur(s):
Koninklijke Philips Electronics N.V.
5621 BA Eindhoven
NETHERLANDS

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Fussschalter für einen Computer

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:

/

Am Anmeldetag benannte Vertragsstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du dépôt:

Bemerkungen:
Remarks:
Remarques:

THIS PAGE BLANK (USPTO)

PHO 99.531 EP-P

- 1 -

Fußschalter für einen Computer

Die Erfindung bezieht sich auf eine Eingabeeinrichtung gemäß dem Oberbegriff des
5 Anspruch 1.

Eine solche mit einem Fuß betätigbare Eingabeeinrichtung gemäß dem Oberbegriff des
Anspruch 1 ist aus dem Dokument WO97/34407 bekannt und durch einen Fußschalter
10 gebildet. Der bekannte Fußschalter weist Anschlußmittel auf, mit denen der Fußschalter
zum Abgeben einer Steuerinformation an einen Game-Port eines Computers anschließbar
ist. An den Computer ist weiters ein Eingangsverstärker, an den ein Mikrofon
angeschlossen ist, ein Ausgangsverstärker, an den ein Kopfhörer angeschlossen ist, eine
Tastatur und ein Monitor anschließbar. Der Computer weist eine A/D-Wandlerstufe, eine
15 D/A-Wandlerstufe und Audio-Speichermittel auf

Mit dem Computer ist ein Anwendungsprogramm abarbeitbar, bei dem eine Diktier-
Betriebsart aktivierbar ist, wobei der Computer eine Diktiereinrichtung bildet. Bei
aktivierter Diktier-Betriebsart wird eine von einem Benutzer in das Mikrofon gesprochene
Sprachinformation eines Diktats als analoges Audiosignal mit dem Eingangsverstärker
20 verstärkt, mit der A/D-Wandlerstufe verarbeitet und als digitale Audioinformation in den
Audio-Speichermitteln des Computers gespeichert.

Bei dem Anwendungsprogramm ist weiters eine Audiowiedergabe-Betriebsart
aktivierbar, bei der die in den Audio-Speichermitteln gespeicherte digitale
Audioinformation von der D/A-Wandlerstufe verarbeitet und als analoges Audiosignal mit
25 dem Ausgangsverstärker verstärkt an den Kopfhörer abgegeben wird. Ein/e Sekretär/in
kann die mit dem Kopfhörer wiedergegebene Sprachinformation des Diktats abhören und
eine entsprechende Textinformation mit der Tastatur des Computers schreiben, worauf die
Textinformation an dem Monitor angezeigt wird.

Der Fußschalter weist mit einem Fuß betätigbare Schalter auf, die Fußeingabemittel
30 bilden. Der/Die Sekretär/in kann durch Betätigen der Schalter mit einem Fuß eine
Steuerinformation erzeugen, die über den Game-Port an den Computer abgebar ist und die
die Audiowiedergabe-Betriebsart aktiviert oder deaktiviert, um die Hände für das

PHO 99.531 EP-P

- 2 -

Schreiben mit der Tastatur frei zu haben.

Bei der bekannten mit einem Fuß betätigbaren Eingabeeinrichtung für einen Computer hat sich als Nachteil erwiesen, daß der Kopfhörer zum Wiedergeben einer in den Audio-Speichermitteln gespeicherten Audioinformation mit einem separaten Kabel an den

5 Computer angeschlossen werden muß, was für den Benutzer auf Grund der vielen Anschlußmöglichkeiten bei einem Computer nicht einfach ist.

Der vorerwähnte Eingangsverstärker und Ausgangsverstärker sind bei einem Computer oft auf einer Sound-Karte vorgesehen, die meistens nur einen Eingangsanschluß für ein Mikrofon und nur einen Ausgangsanschluß für einen Kopfhörer oder einen Lautsprecher

10 aufweist. Hierbei hat sich als Nachteil erwiesen, daß ein Computer nur dann als Diktierereinrichtung verwendbar ist, wenn sowohl eine Sound-Karte mit freien Anschlüssen als auch ein freier Game-Port Anschluß vorhanden sind.

Weiters hat sich als Nachteil erwiesen, daß die Übertragung einer Steuerinformation über den Game-Port des Computers nicht sehr zuverlässig funktioniert und manchmal eine

15 nicht erwünschte Betriebsart des Anwendungsprogramms aktiviert wird.

Die Erfindung hat sich zur Aufgabe gestellt, eine mit einem Fuß betätigbare Eingabeeinrichtung zu schaffen, bei der die Nachteile der bekannten Eingabeeinrichtung

20 nicht auftreten. Diese Aufgabenstellung wird bei einer mit einem Fuß betätigbaren Eingabeeinrichtung gemäß dem Oberbegriff des Anspruch 1 durch die Maßnahmen des kennzeichnenden Teiles des Anspruch 1 gelöst.

Gemäß den Maßnahmen nach Anspruch 1 ist erreicht, daß ein Lautsprecher oder ein Kopfhörer, der Anschlußmittel zum Aufbau einer digitalen Datenbusverbindung aufweist,

25 direkt an die Eingabeeinrichtung anschließbar ist. Mit den Anschlußmitteln des Kopfhörers oder des Lautsprechers empfangene digitale Audiodaten sind von Audio-Verarbeitungsmitteln des Kopfhörers oder des Lautsprechers verarbeitbar und akustisch wiedergebar.

Durch das direkte Anschließen eines Kopfhörers oder eines Lautsprechers zum Abhören

30 eines Diktats an die Eingabeeinrichtung ist sowohl der Vorteil einer einfacheren Installation des Anschlußkabels des Kopfhörers bzw. des Lautsprechers als auch der Vorteil erhalten, daß der Ausgangsanschluß der Sound-Karte und des Game-Ports des

PHO 99.531 EP-P

- 3 -

Computers für andere Anwendungen frei sind. Weiters muß der Computer vorteilhafterweise keine Sound-Karte aufweisen, um als Diktierereinrichtung die Audiowiedergabe-Betriebsart abzuarbeiten.

Bei der Übertragung der Audioinformation und der Steuerinformation über die digitale Datenbusverbindung werden Fehlerkorrekturcodes mit übertragen, weshalb vorteilhafterweise auch bei einer schlechten Kabelverbindung oder einer schlechten Steckverbindung keine Qualitätsverluste der Audioinformation oder der Steuerinformation bei der Übertragung auftreten.

Gemäß den Maßnahmen des Anspruch 2 ist der Vorteil erhalten, daß von den Anschlußmitteln der Eingabeeinrichtung unmittelbar das analoge Audiosignal der Sprachinformation eines in den Audio-Speichermitteln als digitale Audiodaten gespeicherten Diktats an einen herkömmlichen analogen Kopfhörer oder analogen Lautsprecher abgebar ist.

Gemäß den Maßnahmen des Anspruch 3 ist der Vorteil erhalten, daß über eine USB-Datenbusverbindung auch eine Versorgungsspannung und ein ausreichender Versorgungsstrom übertragbar ist, der zur Versorgung der Audio-Verarbeitungsmittel der Eingabeeinrichtung geeignet ist. Weiters weisen bereits sehr viele Computer Anschlußmittel mit einem USB-Anschluß auf, weshalb die Eingabeeinrichtung an viele Computer anschließbar ist.

Gemäß den Maßnahmen des Anspruch 4 ist der Vorteil erhalten, daß eine in den Audio-Speichermitteln gespeicherte Audioinformation als analoges Audiosignal unmittelbar an den in die Eingabeeinrichtung eingebauten Lautsprecher abgebar ist und daher kein weiteres Anschlußkabel nötig ist.

Gemäß den Maßnahmen des Anspruch 5 und des Anspruches 6 ist der Vorteil erhalten, daß keine Kabelverbindung zum Übertragen der Steuerinformation und der Audioinformation notwendig ist.

Die Erfindung wird im Folgenden anhand von drei Anwendungsbeispielen eines in den Figuren dargestellten Ausführungsbeispiels beschrieben, auf das die Erfindung aber nicht beschränkt ist.

Die Figur 1 zeigt eine Spracherkennungseinrichtung in Form eines Blockschaltbilds, an

PHO 99.531 EP-P

- 4 -

die eine in einer Hand haltbare Eingabeeinrichtung und eine mit einem Fuß betätigbare Eingabeeinrichtung angeschlossen sind und die Textvergleichsmittel, zum Vergleichen einer erkannten Textinformation mit einer korrigierten Textinformation, und die Transfermittel, zum Importieren und Exportieren einer Sprachkoeffizienten-Information, aufweist.

Die Figur 2 zeigt die in der Hand haltbare Eingabeeinrichtung gemäß Figur 1 in Form eines Blockschaltbilds, die Audio-Verarbeitungsmittel zum Verarbeiten eines analogen Audiosignals und zum Abgeben digitaler Audiodaten aufweist.

Die Figur 3 zeigt die mit einem Fuß betätigbare Eingabeeinrichtung gemäß Figur 1 in Form eines Blockschaltbilds, an die ein Kopfhörer anschließbar ist und die einen eingebauten Lautsprecher aufweist.

Die Figur 4 zeigt eine von den Textvergleichsmitteln der Spracherkennungseinrichtung erstellte Anpassungstabelle, in der eine erkannte Textinformation und eine korrigierte Textinformation eingetragen sind.

Die Figur 1 zeigt einen Computer 1, mit dem ein Spracherkennungsprogramm gemäß einem Spracherkennungsverfahren abgearbeitet wird, wobei der Computer 1 eine Diktierereinrichtung mit nachgeschalteter Spracherkennungseinrichtung bildet.

An den Computer 1 ist ein Diktiermikrofon 2, das eine in einer Hand haltbare Eingabeeinrichtung bildet, ein Fußschalter 3, der eine mit einem Fuß betätigbare Eingabeeinrichtung bildet, ein Monitor 4 und eine Tastatur 5 angeschlossen. In einer Figur 2 ist das Diktiermikrofon 2 in Form eines Blockschaltbilds detailliert dargestellt.

Das Diktiermikrofon 2 weist ein Mikrofon 6 auf, das Audio-Empfangsmittel bildet und zum Empfangen einer von einem Benutzer in das Mikrofon 6 gesprochenen Sprachinformation und zum Abgeben eines ersten analogen Audiosignals AS1 ausgebildet ist. Das Diktiermikrofon 2 weist weiters Audio-Verarbeitungsmittel 7 auf, die zum Verarbeiten des ersten analogen Audiosignals AS1 und zum Abgeben digitaler Audiodaten als Audioinformation AI ausgebildet sind.

Die Audio-Verarbeitungsmittel 7 weisen eine erste Verstärkerstufe 8, eine A/D-Wandlerstufe 9, eine zweite Verstärkerstufe 10 und eine D/A-Wandlerstufe 11 auf. Das erste analoge Audiosignal AS1 ist der ersten Verstärkerstufe 8 und das verstärkte erste

PHO 99.531 EP-P

- 5 -

analoge Audiosignal ist der A/D-Wandlerstufe 9 zuführbar. Die A/D-Wandlerstufe 9 ist zum Abtasten des verstärkten analogen Audiosignals mit einer Abtastfrequenz von 16kHz und zum Abgeben von Abtastwerten in digitalen Audiodaten ausgebildet, die Datenblöcke mit 16 Bit aufweisen und als Audioinformation AI an USB-Anschlußmittel 12 abgebar
5 sind.

Hierdurch ist der Vorteil erhalten, daß das erste analoge Audiosignal AS1 gleich in dem Diktiermikrofon digitalisiert wird daß die Verstärkerstufe 8 bereits bei der Herstellung des Diktiermikrofons 2 an die maximale Ausgangsspannung des Mikrofons 6 angepaßt werden kann. Es geht somit keine Sprachinformation durch ein Übersteuern der Verstärkerstufe 8
10 verloren, weshalb die digitale Audiodaten enthaltende Audioinformation AI eine gute Qualität aufweist.

Von den USB-Anschlußmitteln 12 ist eine digitale Audiodaten enthaltende Audioinformation AI an die D/A-Wandlerstufe 11 der Audio-Verarbeitungsmittel 7 abgebar. Von der D/A-Wandlerstufe 11 ist ein den digitalen Audiodaten entsprechendes
15 analoges Audiosignal an die zweite Verstärkerstufe 10 abgebar. Ein von der zweiten Verstärkerstufe 10 abgegebenes zweites analoges Audiosignal AS2 ist an einen Lautsprecher 13 zur akustischen Wiedergabe abgebar.

Das Diktiermikrofon 2 weist weiters Eingabemittel 14 zum manuellen Eingeben einer Steuerinformation SI auf, die eine Tastatur 15, einen Trackball 16 und Eingabe-
20 Verarbeitungsmittel 17 enthalten. Die Tastatur 15 weist unter anderem eine Record-Eingabetaste 18, eine Stop-Eingabetaste 19, eine Vorlauf-Eingabetaste 20 und eine Rücklauf-Eingabetaste 21 auf. Von der Tastatur 15 ist eine Tasteninformation TSI an die Eingabe-Verarbeitungsmittel 17 abgebar.

Der Trackball 16 ist durch eine in einer Öffnung des Diktiermikrofons 2 eingelegte
25 Kugel gebildet, die von einem Benutzer gedreht werden kann, um beispielsweise einen mit dem Monitor 4 dargestellten Cursor zu positionieren. Von dem Trackball 16 ist eine Trackballinformation TBI an die Eingabe-Verarbeitungsmittel 17 abgebar. Die Eingabe-Verarbeitungsmittel 17 sind zum Verarbeiten der an sie abgegebenen Tasteninformation TSI und Trackballinformation TBI und zum Abgeben einer entsprechenden
30 Steuerinformation SI an die USB-Anschlußmittel 12 ausgebildet.

Von den USB-Anschlußmitteln 12 ist eine Versorgungsspannung U sowohl an die Audio-Verarbeitungsmittel 7 als auch an die Eingabemittel 14 zur Versorgung der in den

PHO 99.531 EP-P

- 6 -

Audio-Verarbeitungsmitteln 7 und den Eingabemitteln 14 enthaltenen Stufen abgebar. Hierdurch ist der Vorteil erhalten, daß das Diktiermikrofon 2 sowohl Informationen als auch die Versorgungsspannung U der Stufen des Diktiermikrofons 2 über nur eine Kabelverbindung erhält.

- 5 Die USB-Anschlußmittel 12 sind über einen ersten Anschluß 22 des Diktiermikrofons 2 mit USB-Anschlußmitteln 23 des Computers 1 verbunden. Die USB-Anschlußmittel 12 des Diktiermikrofons 2 sind zum Aufbau einer digitalen Datenbusverbindung und hierbei vorteilhafterweise zum Aufbau einer USB-Datenbusverbindung (Universal Serial Bus: Specification Version 1.0 on January 15 and Version 2.0) mit den USB-Anschlußmitteln
- 10 23 des Computers 1 ausgebildet.

- Bei einer USB-Datenbusverbindung werden den an das USB-Anschlußmittel 12 oder 23 abgegebenen Audiodaten der Audioinformation AI oder Steuerdaten der Steuerinformation SI Fehlerkorrekturcodes angehängt und diese gemeinsam als Übertragungsdaten UD an das USB-Anschlußmittel 23 oder 12 übertragen. Das USB-Anschlußmittel 12 oder 23, von
- 15 dem die Übertragungsdaten UD empfangen wurden, ist zum Auswerten der Fehlerkorrekturcodes und zum Korrigieren gegebenenfalls bei der Übertragung aufgetretener Fehler in den Audiodaten oder den Steuerdaten ausgebildet. Bei einer USB-Datenbusverbindung werden Steuerdaten einer Steuerinformation SI mit einer Datenübertragungsrate von 1,5 MBps und Audiodaten einer Audioinformation AI mit einer
- 20 Datenübertragungsrate von 12 MBps übertragen.

- In einer Figur 3 ist der in der Figur 1 dargestellte Fußschalter 3 in Form eines Blockschaltbildes detailliert dargestellt. Der Fußschalter 3 weist Fußeingabemittel 24 zum manuellen Eingeben einer Steuerinformation SI auf. Die Fußeingabemittel 24 enthalten einen Audiowiedergabe-Schalter 25, einen Stop-Schalter 26 und Eingabe-
- 25 Verarbeitungsmittel 27. Eine von dem Audiowiedergabe-Schalter 25 oder dem Stop-Schalter 26 abgegebene Schaltinformation SCI ist an die Eingabe-Verarbeitungsmittel 27 abgebar. Die Eingabe-Verarbeitungsmittel 27 sind zum Verarbeiten der an sie abgegebenen Schaltinformation SCI und zum Abgeben einer entsprechenden Steuerinformation SI an USB-Anschlußmittel 28 des Fußschalters 3 ausgebildet.

- 30 Der Fußschalter 3 weist weiters Audio-Verarbeitungsmittel 29 auf, die eine D/A-Wandlerstufe 30 und eine dritte Verstärkerstufe 31 enthalten. Eine Audiodaten enthaltene Audioinformation AI ist von den USB-Anschlußmitteln 28 an die D/A-Wandlerstufe 30

PHO 99.531 EP-P

- 7 -

abgebbar, welche ein den Audiodaten entsprechendes analoges Audiosignal an die dritte Verstärkerstufe 31 abgibt. Ein von der dritten Verstärkerstufe 31 abgegebenes drittes analoges Audiosignal AS3 ist sowohl an einen in dem Fußschalter 3 vorgesehenen Lautsprecher 32 als auch an einen zweiten Anschluß 33 abgebbar. An den zweiten Anschluß 33 ist ein Kopfhörer 34 angeschlossen. Mit dem Lautsprecher 32 und dem Kopfhörer 34 ist eine akustische Wiedergabe einer Sprachinformation ermöglicht, worauf nachfolgend noch näher eingegangen ist.

Durch das Vorsehen der Audio-Verarbeitungsmittel 29 in dem Fußschalter 3 ist der Vorteil erhalten, daß der Kopfhörer 34 oder ein Lautsprecher direkt an den Fußschalter 3 angeschlossen werden kann. Hierdurch muß der Computer 1, der auch eine Diktierereinrichtung bildet, keine Sound-Karte mit Audio-Verarbeitungsmitteln aufweisen.

Die USB-Anschlußmittel 28 des Fußschalters 3 entsprechen den USB-Anschlußmitteln 12 und 23. Von den USB-Anschlußmitteln 28 ist eine Versorgungsspannung U sowohl an die Eingabe-Verarbeitungsmittel 24 als auch an die Audio-Verarbeitungsmittel 29 zur Versorgung der in den Eingabe-Verarbeitungsmitteln 24 und den Audio-Verarbeitungsmitteln 29 enthaltenen Stufen abgebbar. Hierdurch ist der Vorteil erhalten, daß der Fußschalter 3 sowohl Informationen als auch die Versorgungsspannung U über nur eine Kabelverbindung erhält.

Die USB-Anschlußmittel 28 des Fußschalters 3 sind über einen dritten Anschluß 35 an die USB-Anschlußmitteln 23 des Computers 1 angeschlossen. Der Computer 1 weist Empfangsmittel 36 auf, die die USB-Anschlußmittel 23 und Audio-Speichermittel 37 enthalten und die zum Empfangen einer von einem Sprecher gesprochenen Sprachinformation ausgebildet sind. Hierbei ist eine von einem Sprecher in das Mikrofon 6 des Diktiermikrofons 2 gesprochene Sprachinformation als durch digitale Audiodaten gebildete Audioinformation AI von den USB-Anschlußmitteln 12 an die USB-Anschlußmittel 23 übertragbar und von den USB-Anschlußmitteln 23 an die Audio-Speichermittel 37 zur Speicherung der Audioinformation AI abgebbar. Eine von den USB-Anschlußmitteln 12 des Diktiermikrofons 2 oder von den USB-Anschlußmitteln 28 des Fußschalters 3 an die USB-Anschlußmittel 23 des Computers 1 abgegebene Steuerinformation SI ist von den USB-Anschlußmitteln 23 empfangbar und an weitere Mittel des Computers 1 abgebbar, worauf nachfolgend noch näher eingegangen ist.

Die Spracherkennungseinrichtung, die durch den Computer 1 gebildet ist, weist

PHO 99.531 EP-P

- 8 -

Sprachkoeffizienten-Speichermittel 38 auf, die zum Speichern einer Sprachkoeffizienten-Information SKI ausgebildet sind. Die Sprachkoeffizienten-Information SKI enthält eine Sprachmodell-Information SMI, eine Wortschatz-Information WI und eine Phonemreferenz-Information PRI.

5 Die Wortschatz-Information WI ist in Wortschatz-Speichermitteln 39 gespeichert und enthält sämtliche von der Spracherkennungseinrichtung in einer Sprachinformation erkennbaren Worte sowie einen Häufigkeitszähler je Wort, der die Auftrittswahrscheinlichkeit des Words in einer Sprachinformation angibt. Die Sprachmodell-Information SMI ist in Sprachmodell-Speichermitteln 40 gespeichert und
10 enthält Wortfolgen sowie einen Häufigkeitszähler je Wortfolge, der die Auftrittswahrscheinlichkeiten dieser Wortfolge in einer Sprachinformation angibt. Die Phonemreferenz-Information PRI ist in Phonemreferenz-Speichermitteln 41 gespeichert und ermöglicht ein Anpassen der Spracherkennungseinrichtung an Eigenarten der Aussprache eines Sprechers.

15 Die Spracherkennungseinrichtung weist weiters Spracherkennungsmittel 42 auf. Von den USB-Anschlußmitteln 23 ist eine Steuerinformation SI zur Aktivierung einer Spracherkennungs-Betriebsart zur Abarbeitung eines Spracherkennungsverfahrens an die Spracherkennungsmittel 42 abgebar. Die Spracherkennungsmittel 42 sind bei aktivierter Spracherkennungs-Betriebsart zum Auslesen einer in den Audio-Speichermitteln 37
20 gespeicherten Audioinformation AI und zum Auslesen der in den Sprachkoeffizienten-Speichermitteln 38 gespeicherten Sprachkoeffizienten-Information SKI ausgebildet.

Die Spracherkennungsmittel 42 sind bei der Abarbeitung des Spracherkennungsverfahrens zum Auswerten der Audioinformation AI und der Sprachkoeffizienten-Information SKI ausgebildet. Hierbei ermitteln die
25 Spracherkennungsmittel 42 zu jedem Teil der Audioinformation AI mehrere mögliche Wortfolgen, die gegebenenfalls dem jeweiligen Teil der Audioinformation AI entsprechen, und geben diese möglichen Wortfolgen für alle Teile der Audioinformation AI als mögliche Textinformation PTI ab. Die mögliche Textinformation PTI enthält also beispielsweise 20 mögliche Wortfolgen für die Audioinformation AI. Weiters ermitteln die
30 Spracherkennungsmittel 42 - bei der Abarbeitung des Spracherkennungsverfahrens - die jeweils wahrscheinlichste Wortfolge für jeden Teil der Audioinformation AI und geben diese ermittelten wahrscheinlichsten Wortfolgen der gesamten Audioinformation AI als

PHO 99.531 EP-P

- 9 -

erkannte Textinformation RTI ab.

Die Spracherkennungseinrichtung weist Text-Speichermittel 43 zum Speichern von Textinformationen auf. Die Text-Speichermittel 43 enthalten Möglicher-Text-Speichermittel 44, Erkannter-Text-Speichermittel 45, Korrigierter-Text-Speichermittel 46 sowie Trainings-Text-Speichermittel 47. Eine von den Spracherkennungsmitteln 42 bei der Abarbeitung des Spracherkennungsverfahrens abgegebene mögliche Textinformation PTI ist in den Möglicher-Text-Speichermittel 44 und eine abgegebene erkannte Textinformation RTI ist in den Erkannter-Text-Speichermittel 45 speicherbar.

Der Computer 1 ist weiters zum Abarbeiten eines Textverarbeitungsprogramms – wie beispielsweise „Word for Windows95“ – ausgebildet und bildet hierbei Textverarbeitungsmittel 48. Eine von den Spracherkennungsmitteln 42 bei der Abarbeitung des Spracherkennungsverfahrens abgegebene erkannte Textinformation RTI ist an die Textverarbeitungsmittel 48 abgebar. Die Textverarbeitungsmittel 48 sind hierauf zum Abgeben einer die erkannte Textinformation RTI enthaltenden Bildinformation PI an den Monitor 4 ausgebildet.

Mit den Eingabemitteln 14 des Diktiermikrofons 2 ist eine Steuerinformation SI eingebbar, die an die Spracherkennungsmittel 42 und die Textverarbeitungsmittel 48 abgebar ist und die eine Korrektur-Betriebsart der Spracherkennungseinrichtung aktiviert. Die Textverarbeitungsmittel 48, der Monitor 4 und die Tastatur 5 bilden hierbei Korrekturmittel 49. Bei aktivierter Korrektur-Betriebsart kann ein Benutzer mit der Tastatur 5 die erkannte Textinformation RTI korrigieren und eine korrigierte Textinformation CTI in den Korrigierter-Text-Speichermitteln 46 speichern, worauf nachfolgend noch näher eingegangen ist.

In den Trainings-Text-Speichermitteln 47 ist eine Trainingstextinformation TTI gespeichert, die typische Worte und Redewendungen bzw. Wortfolgen eines bestimmten Bereichs (Geschäftsbrieft; Fachgebiet: Radiologie; Fachgebiet: Recht...) enthält. Mit den Eingabemitteln 14 des Diktiermikrofons 2 ist eine Steuerinformation SI eingebbar, die an die Spracherkennungsmittel 42 und die Textverarbeitungsmittel 48 abgebar ist und die eine Initialtraining-Betriebsart der Spracherkennungseinrichtung aktiviert.

Bei aktivierter Initialtraining-Betriebsart sind die Textverarbeitungsmittel 47 zum Auslesen der Trainingstextinformation TTI aus den Trainingstext-Speichermitteln 47 und zum Abgeben einer entsprechenden Bildinformation PI an den Monitor 4 ausgebildet. Von

PHO 99.531 EP-P

- 10 -

einem Benutzer kann hierauf der mit dem Monitor 4 angezeigte Trainingstext in das Mikrofon 6 gesprochen werden, um die Spracherkennungseinrichtung an die Art der Aussprache des Benutzers anzupassen.

Die Spracherkennungseinrichtung weist nunmehr Anpassungsmittel 50 auf, um die in
5 den Sprachkoeffizienten-Speichermitteln 38 gespeicherte Sprachkoeffizienten-Information SKI sowohl an die Art der Aussprache des Benutzers als auch an von dem Benutzer üblicherweise verwendete Worte und Wortfolgen anzupassen. Die Text-Speichermittel 43, die Korrekturmittel 49, und die Anpassungsmittel 50 bilden hierbei Trainingsmittel 51. Ein solches Anpassen der Sprachkoeffizienten-Information SKI erfolgt bei aktivierter
10 Initialtraining-Betriebsart, bei der die von dem Benutzer vorgelesene Trainingstextinformaion TTI bekannt ist.

Ein solches Anpassen erfolgt aber auch bei einer Anpassungs-Betriebsart, bei der eine einer Sprachinformation entsprechende Textinformation als erkannte Textinformation RTI erkannt und von dem Benutzer in eine korrigierte Textinformation CTI korrigiert wurde.
15 Hierfür weisen die Trainingsmittel 51 Textvergleichsmittel 52 auf, die zum Vergleichen der erkannten Textinformation RTI mit der korrigierten Textinformation CTI und zum Ermitteln zumindest eines Übereinstimmungswerts CI ausgebildet sind. In den Textvergleichsmitteln 52 wird bei aktivierter Anpassungs-Betriebsart eine in einer Figur 4 dargestellte Anpassungstabelle 53 erstellt, worauf nachfolgend noch näher eingegangen ist.

20 Von den Anpassungsmitteln 50 ist als Ergebnis der Anpassung eine Anpassungsinformation NI an die Sprachkoeffizienten-Speichermittel 38 abgebar, mit der die gespeicherte Sprachkoeffizienten-Information SKI angepaßt wird. Von den Anpassungsmitteln 50 der Trainingsmittel 51 ist weiters ein Trainingsindikator TI abgebar, der das Ausmaß der Anpassung der in den Sprachkoeffizienten-Speichermitteln
25 38 gespeicherten Sprachkoeffizienten-Information SKI angibt. Der Trainingsindikator TI enthält hierbei eine Information, wie oft und mit welcher Anzahl von Worten die in den Sprachkoeffizienten-Speichermitteln 38 gespeicherte Sprachkoeffizienten-Information SKI bereits an einen Benutzer angepaßt wurde. Je öfter die Sprachkoeffizienten-Information SKI an einen Benutzer angepaßt wurde, desto besser ist die Erkennungsrate der
30 Spracherkennungseinrichtung für diesen Benutzer.

Die Spracherkennungseinrichtung weist nunmehr Transfermittel 54 auf, die ein Importieren einer Sprachkoeffizienten-Information SKI und ein Speichern der importierten

PHO 99.531 EP-P

- 11 -

Sprachkoeffizienten-Information SKI in den Sprachkoeffizienten-Speichermitteln 38 und/oder ein Exportieren der in den Sprachkoeffizienten-Speichermitteln 38 gespeicherten Sprachkoeffizienten-Information SKI ermöglichen. Hierfür ist von dem Benutzer eine Transferinformation TRI mit der Tastatur 5 eingebbar und an die Transfermittel 54
5 abgebbbar, worauf nachfolgend näher eingegangen ist.

Anhand eines ersten Anwendungsbeispiels ist nunmehr die Funktion des Diktiermikrofons 2, des Fußschalters 3 und der durch den Computer 1 gebildeten Diktiereinrichtung und Spracherkennungseinrichtung näher erläutert. Gemäß dem ersten Anwendungsbeispiel ist angenommen, daß der Benutzer des Computers 1 einen Brief
10 diktieren möchte, um bei seinem Bekannten „John“ drei Computer des Typs „Quality“ zu bestellen. Hierfür betätigt der Benutzer die Record-Eingabetaste 18 des Diktiermikrofons 2 und diktiert „Dear John I herewith order three computer Quality for my office Harry“. Eine dieser Sprachinformation entsprechend Audioinformation AI wird hierauf von dem Diktiermikrofon 2 als Übertragungsdaten UD an die USB-Anschlußmittel 23 übertragen
15 und schließlich in den Audio-Speichermitteln 37 gespeichert. Bei diesem Vorgang bilden das Diktiermikrofon 2 und der Computer 1 eine Diktiereinrichtung.

Durch das Betätigen der Record-Eingabetaste 18 des Diktiermikrofons 2 wird eine die Spracherkennungs-Betriebsart der Spracherkennungsmittel 42 aktivierende Steuerinformation SI von dem Diktiermikrofon 2 als Übertragungsdaten UD über die USB-
20 Anschlußmittel 23 an die Spracherkennungsmittel 42 abgegeben. Bei in den Spracherkennungsmitteln 42 aktivierter Spracherkennungs-Betriebsart wird die in den Audio-Speichermitteln 37 gespeicherte Audioinformation AI ausgelesen und gemeinsam mit der aus den Sprachkoeffizienten-Speichermitteln 38 ausgelesenen Sprachkoeffizienten-Information SKI ausgewertet. Hierbei ermitteln die Spracherkennungsmittel 42 mehrere
25 mögliche Wortfolgen als mögliche Textinformation PTI und speichern diese in den Möglicher-Text-Speichermitteln 44. Eine mögliche Textinformation PTI könnte hierbei für den ersten Teil der Audioinformation AI beispielsweise die Wortfolgen „Hear John why“, „Dear John I“ und „Bear John hi“ enthalten.

Die Spracherkennungsmittel 42 ermitteln die Textinformation „Dear John I herewith
30 organ three computer Quality for my office Mary“ als wahrscheinlichste Textinformation und geben diese als erkannte Textinformation RTI an die Erkannter-Text-Speichermittel 45 ab. Hierbei wurde das gesprochene Wort „order“ als das Wort „organ“ und das

PHO 99.531 EP-P

- 12 -

gesprochene Wort „Harry“ als das Wort „Mary“ falsch erkannt.

Die von den Spracherkennungsmittel 42 erkannte und in den Erkannter-Text-Speichermitteln 45 gespeicherte erkannte Textinformation RTI wird hierauf mit den Textverarbeitungsmitteln 48 ausgelesen und mit dem Monitor 4 dargestellt. Der Benutzer
5 erkennt, daß die beiden gesprochenen Worte „order“ und „Harry“ falsch erkannt wurden und möchte die erkannte Textinformation RTI korrigieren, weshalb der Benutzer mit den Eingabemitteln 14 des Diktiermikrofons 2 die Korrektur-Betriebsart der Spracherkennungseinrichtung aktiviert.

Bei aktivierter Korrektur-Betriebsart kann der Benutzer durch Betätigen des
10 Audiowiedergabe-Schalters 25 des Fußschalters 3 die in den Audio-Speichermitteln 37 gespeicherte Audioinformation AI mit dem Kopfhörer 34 oder dem Lautsprecher 32 akustisch wiedergeben, wobei synchron zu der wiedergegebenen Audioinformation AI die zu der Audioinformation AI erkannte Textinformation RTI mit dem Monitor 4 gekennzeichnet wird. Ein solches synchrones Editieren bzw. Korrigieren ist seit langem
15 bekannt.

Bei einer ersten Art der Korrektur der erkannten Textinformation RTI durch den Benutzer korrigiert der Benutzer nur die falsch erkannten Worte „organ“ und „Mary“ in die beiden eigentlich gesprochenen Worte „order“ und „Harry“ und es wird eine korrigierte Textinformation CTI in den Textverarbeitungsmitteln 48 erstellt und in den Korrigierter-
20 Text-Speichermitteln 46 gespeichert. Diese nach der ersten Art der Korrektur in den Korrigierter-Text-Speichermitteln 46 gespeicherte korrigierte Textinformation CTI wäre zur Anpassung der Sprachkoeffizienten-Information SKI besonders gut geeignet. Die bei einem weiteren Diktat des Benutzers gesprochenen Worte „order“ und „Harry“ würden bei einem darauffolgenden Spracherkennungsvorgang mit der zuvor angepaßten
25 Sprachkoeffizienten-Information SKI richtig erkannt werden.

Wie die Praxis zeigt, werden von Benutzern bei aktivierter Korrektur-Betriebsart auch andere Korrekturen gemäß einer zweiten Art der Korrektur durchgeführt. Gemäß dem ersten Anwendungsbeispiel ist angenommen, daß sich der Benutzer nunmehr erinnert, daß er ebenfalls drei Computer des Typs „Standard“ bestellen möchte, weshalb der Benutzer
30 die Wortfolge „and three computer Standard“ mit der Tastatur 5 in die bereits korrigierte Textinformation CTI einfügt. Weiters kopiert der Benutzer einen Text A vor den Anfang der bereits korrigierten Textinformation CTI, der die Adresse des Briefes an seinen

PHO 99.531 EP-P

- 13 -

Bekannten „John“ enthält. Letztlich kopiert der Benutzer einen Text B nach der bereits korrigierten Textinformation CTI in die korrigierte Textinformation CTI hinein, die eine bei solchen Bestellungen übliche Fußnote über vereinbarte Prozentnachlässe beinhaltet.

Hierbei wurden von dem Benutzer Textteilinformationen in die erkannte

- 5 Textinformation RTI eingefügt, die mit der von dem Benutzer bei dem Diktat gesprochenen Sprachinformation keinen Zusammenhang haben. Solche Textteilinformationen dürfen nicht zur Anpassung der Phonemreferenz-Information PRI ausgewertet werden, da keine entsprechende Audioinformation AI in den Audio-Speichermitteln 37 gespeichert ist. Weiters ist es nicht gut, die Wortschatz-Information WI
- 10 oder die Sprachmodell-Information SMI mit solchen Textteilinformationen anzupassen, da der Text A und der Text B oft in korrigierten Textinformationen CTI des Benutzers enthalten sind, in Diktaten des Benutzers aber praktisch nie gesprochen werden.

- Wenn die Wortschatz-Information WI und die Sprachmodell-Information SMI mit Textteilinformationen trainiert werden würden, die keinen Zusammenhang mit der
- 15 Sprachinformation aufweisen, dann hätte beispielsweise die Wortfolge des in dem Text A enthaltenen Straßennamen „Hellroad“ der Adresse des Bekannten John eine hohe Auftrittswahrscheinlichkeit, da der Benutzer oft Briefe an John schreibt und die Adresse immer in die korrigierte Textinformation CTI hinein kopiert. Nunmehr würde aber eine in einem Diktat gesprochene Wortfolge „Hello Rod“ relativ leicht falsch erkannt werden, da
- 20 die Wortfolge „Hello Rod“ in den Sprachmodell-Speichermitteln 40 eine verglichen mit der Wortfolge „Hellroad“ relativ niedrigere Auftrittswahrscheinlichkeit aufweisen würde, obwohl die Wortfolge „Hello Rod“ in Diktaten des Benutzers relativ häufig vorkommt.

- Nachdem der Benutzer alle von ihm gewünschten Korrekturen durchgeführt hat und die korrigierte Textinformation CTI in den Korrigierte-Text-Speichermitteln 46 gespeichert
- 25 wurde, betätigt der Benutzer eine Taste der Tastatur 15 des Diktiermikrofons 2, um die Anpassungs-Betriebsart der Spracherkennungseinrichtung zu aktivieren und die Sprachkoeffizienten-Information SKI der Spracherkennungseinrichtung anzupassen. Eine entsprechende Steuerinformation SI wird hierauf von dem Diktiermikrofon 2 über die USB-Anschlußmittel 23 an die Spracherkennungsmittel 42 abgegeben.

- 30 Bei in der Spracherkennungseinrichtung aktivierter Anpassungs-Betriebsart sind die Anpassungsmittel 50 zum Auslesen der in den Audio-Speichermitteln 37 gespeicherten Audioinformation AI, sowie der in den Text-Speichermitteln 43 gespeicherten möglichen

PHO 99.531 EP-P

- 14 -

Textinformation PTI, erkannten Textinformation RTI und korrigierten Textinformation CTI ausgebildet. Weiters sind die Textvergleichsmittel 52 bei aktivierter Anpassungs-Betriebsart zum Erstellen der in der Figur 4 dargestellten Anpassungstabelle 53 ausgebildet, um jene Textteilinformationen der korrigierten Textinformation CTI zu
5 ermitteln, die zur Anpassung der Sprachkoeffizienten-Information SKI geeignet sind.

Hierfür wird vorerst die erkannte Textinformation RTI und die korrigierte Textinformation CTI in die Anpassungstabelle 53 eingetragen, wobei falsch erkannte Worte sowie korrigierte und eingefügte Worte und Textteile mit der Schriftart „fett“ dargestellt sind. Die Textvergleichsmittel 52 sind nunmehr zum Vergleichen der erkannten
10 Textinformation RTI und der korrigierten Textinformation CTI und zum Kennzeichnen übereinstimmender Worte durch graue Felder an der entsprechenden Stelle der Anpassungstabelle 53 ausgebildet. Hierbei wird die Textteilinformation mit den Worten „three computer“ der erkannten Textinformation RTI zwei Mal in der korrigierten Textinformation CTI gefunden.

15 Weiters sind die Textvergleichsmittel 52 beim Vergleichen der erkannten Textinformation RTI und der korrigierten Textinformation CTI zum Ermitteln eines Übereinstimmungswerts CI für jede Textteilinformation ausgebildet. Hierbei ermitteln die Textvergleichsmittel 52 wie viele übereinstimmende durch ein graues Feld gekennzeichnete Worte eine Textteilinformation aufweist. Weiters ermitteln die
20 Textvergleichsmittel 52 Strafpunkte für jede Textteilinformation, wobei für jede Einfügung, Löschung oder Ersetzung eines Wortes in der korrigierten Textinformation CTI ein Strafpunkt vergeben wird. Aus der Anzahl der übereinstimmenden Worte und der Strafpunkte einer Textteilinformation wird der Übereinstimmungswert CI der Textteilinformation ermittelt.

25 In den Textvergleichsmitteln 52 ist ein Mindestwert MW für den Übereinstimmungswert CI festgelegt, der unterschritten wird, wenn für eine Textteilinformation mehr als drei Strafpunkte für Korrekturen benachbarter Worte der korrigierten Textinformation CTI vergeben werden. Für die Anpassung der Sprachkoeffizienten-Information SKI werden nur Textteilinformationen verwendet, deren
30 Übereinstimmungswert CI den Mindestwert MW überschreitet.

Die Textvergleichsmittel 52 werten nunmehr die Anpassungstabelle 53 aus und erkennen, daß sämtliche Worte des Text A und des Text B in der korrigierten

PHO 99.531 EP-P

- 15 -

Textinformation CTI eingefügt wurden, weshalb die Textvergleichsmittel 52 für den Text A und den Text B eine mit der Anzahl der Worte des Text A und das Text B übereinstimmende Anzahl an Strafpunkten vergeben. Die Textvergleichsmittel 52 geben daher einen Übereinstimmungswert CI für den Text A und den Text B an die

- 5 Anpassungsmittel 50 ab, der kleiner als der Mindestwert MW ist. Hierdurch ist der Vorteil erhalten, daß eine Textteilinformation der korrigierten Textinformation CTI, die überhaupt keinen Zusammenhang mit der Sprachinformation aufweist, für die Anpassung der Sprachkoeffizienten-Information SKI nicht verwendet wird.

- 10 Gemäß der Anpassungstabelle 53 verbleiben nunmehr Textteilinformationen entlang eines Pfades P1 und eines Pfades P2 der korrigierten Textinformation CTI, die für eine Anpassung der Sprachkoeffizienten-Information SKI verwendet werden könnten.

- Der Pfad P1 enthält eine erste Textteilinformation „Dear John I herewith“, der von den Textvergleichsmitteln 52 keine Strafpunkte zugeordnet werden und deren Übereinstimmungswert daher den Mindestwert MW überschreitet. Der Pfad P1 enthält
- 15 nach der ersten Textteilinformation eine zweite Textteilinformation „order three computer Quality and“ der fünf Strafpunkte (1) bis (5) für die Einfügung der Worte „order three computer Quality“ und die Ersetzung des Wortes „organ“ durch das Wort „and“ zugeordnet werden und deren Übereinstimmungswert daher den Mindestwert MW unterschreitet. Der Pfad P1 enthält nach der zweiten Textteilinformation eine dritte
- 20 Textteilinformation „three computer Standard for my office Harry“, der nur ein Strafpunkt (6) für die Einfügung des Wortes „Standard“ zugeordnet wird und deren Übereinstimmungswert CI daher den Mindestwert MW überschreitet. Der Pfad P1 enthält daher die erste und die dritte Textteilinformation, die prinzipiell für eine Anpassung der Sprachkoeffizienten-Information SKI geeignet sind.

- 25 Der Pfad P2 enthält eine vierte Textteilinformation „Dear John I herewith order three computer Quality“, der von den Textvergleichsmitteln 52 nur ein Strafpunkt (1) für die Ersetzung des Wortes „organ“ durch das Wort „order“ zugeordnet wird und deren Übereinstimmungswert CI daher den Mindestwert MW überschreitet. Der Pfad P2 enthält nach der vierten Textteilinformation eine fünfte Textteilinformation „and three computer
- 30 Standard“ der vier Strafpunkte (2) bis (5) für die Einfügung dieser Worte zugeordnet wird und deren Übereinstimmungswert daher den Mindestwert MW unterschreitet. Der Pfad P2 enthält nach der fünften Textteilinformation eine sechste Textteilinformation „for my

PHO 99.531 EP-P

- 16 -

office Harry“, der ein Strafpunkt (6) für die Ersetzung des Wortes „Mary“ durch das Wort „Harry“ zugeordnet wird und deren Übereinstimmungswert CI den Mindestwert MW überschreitet. Der Pfad P2 enthält daher die vierte und die sechste Textteilinformation, die prinzipiell für eine Anpassung der Sprachkoeffizienten-Information SKI geeignet sind.

- 5 Die Textvergleichsmittel 52 sind nunmehr zum Ermitteln der Textteilinformationen des Pfades P1 oder P2 ausgebildet, die den größten Summen-Übereinstimmungswert SCI aufweisen. Hierfür ermitteln die Textvergleichsmittel 52, daß neun Worte des Pfades P1 und zehn Worte des Pfades P2 der korrigierten Textinformation CTI vollkommen mit Worten der erkannten Textinformation RTI übereinstimmen und in der Anpassungstabelle
- 10 53 als graues Feld gekennzeichnet sind. Weiters werden von den Textvergleichsmitteln 52 die Strafpunkte jedes Pfades addieren, wobei für den Pfad P1 sieben Strafpunkte und für den Pfad P2 sechs Strafpunkte ermittelt werden.

- Schließlich ermitteln die Textvergleichsmittel 52, daß die Textinformationen des Pfades P2 mit zehn übereinstimmenden Worten und nur sechs Strafpunkten eine größere
- 15 Summen-Übereinstimmung mit der erkannten Textinformation RTI aufweisen und geben einen entsprechend hohen Summen-Übereinstimmungswert SCI für die Textteilinformationen des Pfades P2 an die Anpassungsmittel 50 ab.

- Hierdurch ist der Vorteil erhalten, daß für die Anpassung der Sprachkoeffizienten-Information SKI gemäß der ersten Art der Korrektur korrigierte Worte - wie beispielsweise
- 20 „order“ und „Harry“ - zur Anpassung verwendet werden und gemäß der zweiten Art der Korrektur korrigierte Textteilinformationen - wie beispielsweise „and three computer Standard“, die mit der Sprachinformation keinen Zusammenhang aufweisen – nicht zur Anpassung verwendet werden. Hierdurch wird die Sprachkoeffizienten-Information SKI besonders rasch und gut an Eigenarten der Aussprache eines Benutzers und typische
- 25 Formulierungen eines Benutzers angepaßt, womit eine besonders gute Erkennungsrate der Spracherkennungseinrichtung erzielt wird.

- Die Anpassungsmittel 50 sind bei in der Spracherkennungseinrichtung aktivierter Anpassungs-Betriebsart nach dem Erhalt des Summen-Übereinstimmungswerts SCI von den Textvergleichsmitteln 52 zum Auslesen der Textteilinformationen des Pfades P2 der
- 30 korrigierten Textinformation CTI aus den Korrigierter-Text-Speichermitteln 46 und zum Auswerten dieser Textteilinformationen ausgebildet.

Hierbei werden zur Anpassung der Wortschatz-Information WI jedem Wort der vierten

PHO 99.531 EP-P

- 17 -

und der sechsten Textteilinformation „Dear“, „John“, „I“, ..., „Harry“ der Wortfolge P2 in den Wortschatz-Speichermitteln 39 zugeordnet gespeicherten Häufigkeitszähler der Wert „1“ hinzu addiert, wobei der Häufigkeitszähler die Auftrittswahrscheinlichkeit des jeweiligen Worts in einem Diktat des Benutzers kennzeichnet. Wenn ein Wort der zur

5 Anpassung verwendeten Textteilinformation erstmals von dem Benutzer diktiert wurde, dann wird es in den Wortschatz-Speichermitteln 39 gemeinsam mit einem Häufigkeitszähler mit dem Wert „1“ gespeichert.

Zur Anpassung der in den Sprachmodell-Speichermitteln 40 gespeicherten Sprachmodell-Information SMI wird beispielsweise der der Wortfolge „I herewith order“

10 zugeordnet gespeicherte Häufigkeitszähler um den Wert „1“ erhöht, um eine höhere Wahrscheinlichkeit des Auftretens einer solchen Formulierung in einem Diktat des Benutzers zu kennzeichnen.

Zur Anpassung der in den Phonemreferenz-Speichermitteln 41 gespeicherten Phonemreferenz-Information PRI werden nur solche Worte der vierten und der sechsten

15 Textteilinformation der Wortfolge P2 gemeinsam mit der zugehörigen Audioinformation AI ausgewertet, die eine völlige Übereinstimmung mit einem Wort der erkannten Textinformation RTI aufweisen und in der Anpassungstabelle 53 durch ein graues Feld gekennzeichnet sind. Zusätzlich können auch noch Worte der möglichen Textinformation PTI mit der zugehörigen Audioinformation AI ausgewertet werden. Hierfür prüfen die

20 Anpassungsmittel 50 ob ein von dem Benutzer bei der ersten Art der Korrektur ersetztes Wort („order“, „Harry“) in den Möglicher-Text-Speichermitteln 44 für diesen Teil der Audioinformation AI gespeichert ist. Wenn ein solches Wort in den Möglicher-Text-Speichermitteln 44 gefunden wird, dann wird auch diese Wort mit der zugehörigen Audioinformation AI zur Anpassung der Phonemreferenz-Information PRI ausgewertet.

25 Eine diese Anpassungen enthaltende Anpassungsinformation NI wird anschließend von den Anpassungsmitteln 50 an die Sprachkoeffizienten-Speichermittel 38 zur Anpassung der Sprachkoeffizienten-Information SKI abgegeben. Ebenso wird von den Anpassungsmitteln 50 ein Trainingsindikator TI an die Transfermittel 54 abgegeben, der kennzeichnet, mit wie vielen Worten die in den Sprachkoeffizienten-Speichermitteln 38

30 gespeicherte Sprachkoeffizienten-Information SKI bereits trainiert wurde. Hierdurch ist der

Vorteil erhalten, daß anhand des Trainingsindikators TI unmittelbar ersichtlich ist, ob die Spracherkennungseinrichtung bereits gut an einen Benutzer angepaßt wurde und daher

PHO 99.531 EP-P

- 18 -

voraussichtlich eine gute Erkennungsrate für diesen Benutzer aufweisen wird.

Gemäß einem zweiten Anwendungsbeispiel ist angenommen, daß der Benutzer der durch den Computer 1 gebildeten Spracherkennungseinrichtung auch einen anderen Computer besitzt, der eine andere Spracherkennungseinrichtung bildet. Der Benutzer
5 möchte verständlicherweise bei der anderen Spracherkennungseinrichtung eine gleich gute Erkennungsrate erzielen, wie er sie bei der durch den Computer 1 gebildeten Spracherkennungseinrichtung bereits erzielt, deren Sprachkoeffizienten-Information SKI er seit mehreren Monaten trainiert hat.

Hierfür betätigt der Benutzer eine Taste der Tastatur 5, worauf eine Transferinformation
10 TRI von der Tastatur 5 an die Transfermittel 54 angegeben wird. Die Transfermittel 54 sind hierauf zum Auslesen der in den Sprachkoeffizienten-Speichermitteln 38 gespeicherten Sprachkoeffizienten-Information SKI ausgebildet. Weiters sind die Transfermittel 54 zum Speichern der ausgelesenen Sprachkoeffizienten-Information SKI und des Trainingsindikators TI der Sprachkoeffizienten-Information SKI auf einer in ein
15 Diskettenlaufwerk des Computers 1 eingelegten Diskette 55 ausgebildet.

Hierdurch ist der Vorteil erhalten, daß eine in den Sprachkoeffizienten-Speichermittel 38 gespeicherte und bereits von dem Benutzer trainierte Sprachkoeffizienten-Information SKI von dem Computer 1 exportiert werden kann. Weiters kann die Sprachkoeffizienten-Information SKI in dem anderen Computer des Benutzers importiert werden. Hierdurch
20 kann ein Benutzer eine von ihm trainierte Sprachkoeffizienten-Information SKI bei beliebig vielen anderen Computern, die eine der in der Figur 1 dargestellten Spracherkennungseinrichtung entsprechende Spracherkennungseinrichtung bilden, weiter verwenden und so auch bei diesen Computern eine gute Erkennungsrate der Spracherkennungseinrichtung erhalten.

Vorteilhafterweise wird beim Exportieren der Sprachkoeffizienten-Information SKI
25 ebenso der Trainingsindikator TI der Sprachkoeffizienten-Information SKI mit exportiert, womit bei einem Importieren der Sprachkoeffizienten-Information SKI von den Transfermitteln 54 vorerst prüfbar ist, ob die zu importierende Sprachkoeffizienten-Information SKI bereits besser trainiert ist als die in den Sprachkoeffizienten-

30 Speichermitteln 38 gespeicherte Sprachkoeffizienten-Information SKI. Die Transfermittel 54 importieren die Sprachkoeffizienten-Information SKI nur dann, wenn die Prüfung ergibt, daß die zu importierende Sprachkoeffizienten-Information SKI bereits besser

PHO 99.531 EP-P

- 19 -

trainiert ist als die in den Sprachkoeffizienten-Speichermitteln 38 gespeicherte Sprachkoeffizienten-Information SKI.

Gemäß einem dritten Anwendungsbeispiel ist angenommen, daß der Computer 1 einen mit dem Internet verbundenen Internet-Server bildet und mit einem vierten Anschluß 56 mit dem Internet verbunden ist. Weiters sei angenommen, daß der Benutzer eines weiteren Computers, der eine Diktiereinrichtung bildet, die in Homepage-Speichermitteln 57 des Computers 1 gespeicherte Textinformation einer Homepage des Internet-Servers abgefragt hat und diese Textinformation über den vierten Anschluß 56 an den weiteren Computer abgegeben wurde. Diese Textinformation der Homepage enthält Link-Informationen, mit denen weitere in den Homepage-Speichermitteln 57 gespeicherte Textinformationen abgefragt werden können, wie dies allgemein üblich ist. Die in der Figur 1 dargestellte Spracherkennungseinrichtung des Internet-Servers ermöglicht dem Benutzer des weiteren Computers ein Aktivieren einer in der Homepage dargestellten Link-Information durch eine mit dem weiteren Computer eingegebene Sprachinformation.

Gemäß dem dritten Anwendungsbeispiel ist weiters angenommen, daß der Benutzer die in der Homepage angegebene Link-Information „Hotels“ aktivieren möchte. Hierfür betätigt der Benutzer eine Record-Eingabetaste eines Diktiermikrofons des weiteren Computers und spricht die Sprachinformation „Hotels“ in das Mikrofon des Diktiermikrofons. Eine diese Sprachinformation enthaltende Audioinformation AI wird dem Computer 1 über den vierten Anschluß 56 zugeführt und die Audioinformation AI wird in den Audio-Speichermitteln 37 gespeichert. Bei der hierdurch aktivierten Spracherkennungs-Betriebsart der Spracherkennungseinrichtung des Computers 1 wird das Wort „Hotels“ von den Spracherkennungsmitteln 42 als erkannte Textinformation RTI erkannt und die dieser Link-Information zugeordnete Textinformation wird von den Homepage-Speichermitteln 57 über den vierten Anschluß 56 an den weiteren Computer des Benutzers abgegeben.

Bei dem Computer 1 wird nunmehr vorteilhafterweise vor der Aktivierung der Spracherkennungs-Betriebsart die von dem weiteren Computer exportierte und über das Internet an den vierten Anschluß 56 abgegebene Sprachkoeffizienten-Information SKI mit den Transfermitteln 54 importiert. Hierdurch wird der durch den Computer 1 gebildete Internet-Server an den jeweiligen Benutzer angepaßt, weshalb eine besonders hohe Erkennungsrate erzielt wird.

Es kann erwähnt werden, daß die Sprachkoeffizienten-Information SKI eines Benutzers

PHO 99.531 EP-P

- 20 -

jeweils gemeinsam mit jeder an den Internet-Server übertragen Audioinformation AI übertragen werden kann. Besonders vorteilhaft ist es jedoch, die Sprachkoeffizienten-Information SKI eines Benutzers nur bei einem ersten Abfragen der Homepage des Internet-Servers an den Internet-Server zu übertragen und die Sprachkoeffizienten-Information SKI in dem Internet-Server für diesen Benutzer zu speichern. Bei jedem weiteren Empfang einer Audioinformation AI von dem weiteren Computer dieses Benutzers kann der Internet-Server die bereits gespeicherte Sprachkoeffizienten-Information SKI für die Spracherkennung verwenden.

Es kann erwähnt werden, daß die Anpassungs-Betriebsart zum Anpassen der Sprachkoeffizienten-Information SKI auch automatisch nach dem Abspeichern einer korrigierten Textinformation CTI oder durch den Benutzer mit einer Taste der Tastatur 5 des Computers 1 aktiviert werden kann.

Es kann erwähnt werden, daß es auch andere als gemäß dem Ausführungsbeispiel beschriebene Möglichkeiten der Ermittlung eines Übereinstimmungswertes einer Textteilinformation oder eines Summen-Übereinstimmungswertes von Textteilinformationen eines Pfades gibt. Weiters ist die Festlegung, daß der Mindestwert MW unterschritten ist, wenn drei aufeinanderfolgende Strafpunkte vergeben werden, nur als Beispiel zu verstehen. Ebenso könnte eine bestimmte Anzahl an Strafpunkten in einer bestimmten Anzahl an Worten einer Textteilinformation überschreitende Anzahl an Strafpunkten als ein Unterschreiten des Mindestwertes MW festgelegt werden.

Es kann erwähnt werden, daß es auch vorteilhaft sein kann aus einer prinzipiell zur Anpassung der Sprachkoeffizienten-Information SKI geeigneten Textteilinformation, deren Ermittlung anhand des ersten Anwendungsbeispiels erläutert wurde, nur bestimmte Teile zur Anpassung der Sprachkoeffizienten-Information SKI zu verwenden. Hierbei können beispielsweise nur Worte, denen von den Textvergleichsmitteln 52 ein Strafpunkt zugeordnet wurde, und diesen Worten benachbarte Worte, die in der Anpassungstabelle 53 durch ein graues Feld gekennzeichnet sind, verwendet werden.

Hierdurch ist der Vorteil erhalten, daß bei dem Spracherkennungsvorgang falsch erkannte und von dem Benutzer korrigierte (ersetzte) Worte („order“, „Harry“) sowie zu diesen Worten benachbarte Worte zur Anpassung verwendet werden. Teile der korrigierten Textteilinformation CTI von prinzipiell zur Anpassung geeigneten Textteilinformationen, die vollkommen mit der erkannten Textinformation RTI übereinstimmen, werden hierbei

PHO 99.531 EP-P

- 21 -

nicht zur Anpassung verwendet, da diese Teile eventuell von dem Benutzer gar nicht korrigiert wurden.

Es kann erwähnt werden, daß Eingabemittel einer in einer Hand haltbaren Eingabeeinrichtung auch zum Abgeben einer Steuerinformation an einen Computer ausgebildet sein können, um den Computer aus einer Power-Save-Betriebsart in eine Aktiv-Betriebsart aufzuwecken.

Es kann erwähnt werden, daß eine mit einem Fuß betätigbare Eingabeeinrichtung auch ganz allgemein zur Aktivierung von Betriebszuständen in dem Computer verwendbar ist. Hierbei könnte der Computer einen Fernsehempfänger bilden und mit dem Fußschalter beispielsweise ein Fernsehkanal ausgewählt werden, in dem eine Fernsehsendung empfangen wird, die mit dem Monitor wiedergegeben wird.

Es kann erwähnt werden, daß an eine mit einem Fuß betätigbare Eingabeeinrichtung auch ein Mikrofon oder ein Lautsprecher an den USB-Anschlußmitteln der Eingabeeinrichtung angeschlossen werden kann, wenn das Mikrofon oder der Lautsprecher ebenfalls USB-Anschlußmittel aufweist. In diesem Fall ist es nicht notwendig, daß die mit einem Fuß betätigbare Eingabeeinrichtung Audio-Verarbeitungsmittel aufweist, da Audio-Verarbeitungsmittel zum Abgeben eines analogen Audiosignals dann in dem Mikrofon und dem Lautsprecher vorgesehen sind.

Es kann erwähnt werden, daß eine digitale Datenbusverbindung auch durch eine Radio Frequency Funkverbindung, wie beispielsweise entsprechend der Bluetooth-Spezifikation (Bluetooth Special Interest Group, May 1998), gebildet sein kann. Hierdurch ergibt sich der Vorteil, daß keine Kabelverbindung zwischen der Eingabeeinrichtung und einem Computer notwendig sind.

Es kann erwähnt werden, daß an den zweiten Anschluß 33 des Fußschalters 3 auch ein Lautsprecher angeschlossen werden kann.

THIS PAGE BLANK (USPTO)

PHO 99.531 EP-P

- 22 -

Patentansprüche:

1. Mit einem Fuß betätigbare Eingabeeinrichtung (3) für einen Computer (1), der eine Diktiereinrichtung bildet, mit
Fußeingabemitteln (24) zum manuellen Eingeben einer Steuerinformation (SI), mit der
5 eine Audiowiedergabe-Betriebsart der Diktiereinrichtung aktivierbar (25) und deaktivierbar (26) ist, wobei bei aktivierter Audiowiedergabe-Betriebsart eine in der Diktiereinrichtung gespeicherte Audioinformation (AI) an einen Kopfhörer (34) bzw. einen Lautsprecher (32) abgebar ist, und mit
Anschlußmitteln (28, 33, 35) zum Anschließen der Eingabeeinrichtung (3) an den
10 Computer (1), wobei von den Anschlußmitteln (28, 33, 35) die Steuerinformation (SI) an den Computer (1) abgebar ist,
dadurch gekennzeichnet,
daß an die Anschlußmittel (28, 33, 35) eine digitale Datenbusverbindung anschließbar ist und daß die Anschlußmittel zum Empfangen der Audioinformation (AI) als digitale
15 Audiodaten von dem Computer (1) und zum Abgeben der Steuerinformation (SI) als Steuerdaten an den Computer (1) über die digitale Datenbusverbindung ausgebildet sind und
daß an die Anschlußmittel (28, 33, 35) der Kopfhörer (34) bzw. der Lautsprecher anschließbar ist.
- 20 2. Eingabeeinrichtung (3) nach Anspruch 1, dadurch gekennzeichnet, daß Audio-Verarbeitungsmittel (29) vorgesehen sind, die zum Verarbeiten der mit den Anschlußmitteln (28, 33, 35) empfangenen Audioinformation (AI) und zum Abgeben eines analogen Audiosignals (AS3) für den Kopfhörer (34) bzw. den Lautsprecher an die Anschlußmittel (33) ausgebildet sind.
- 25 3. Eingabeeinrichtung (3) nach Anspruch 2, dadurch gekennzeichnet, daß die Anschlußmittel (28) zum Aufbau einer USB-Datenbusverbindung (Universal Serial Bus: Specification Version 1.0 on January 15 or Version 2.0) ausgebildet sind.
4. Eingabeeinrichtung (3) nach Anspruch 1, dadurch gekennzeichnet, daß ein Lautsprecher (32) vorgesehen ist und daß die Audio-Verarbeitungsmittel (29) zum
30 Abgeben des analogen Audiosignals an den eingebauten Lautsprecher (32) ausgebildet sind.
5. Eingabeeinrichtung (3) nach Anspruch 1, dadurch gekennzeichnet, daß

PHO 99.531 EP-P

- 23 -

Anschlußmittel zum Aufbau einer digitalen Datenbusverbindung über eine Radio Frequency Funkverbindung ausgebildet sind.

6. Eingabeeinrichtung nach Anspruch 5, dadurch gekennzeichnet, daß die Anschlußmittel zum Aufbau einer der Bluetooth-Spezifikation (Bluetooth Special Interest Group, May 1998) entsprechenden Radio Frequency Funkverbindung ausgebildet sind.
- 5

PHO 99.531 EP-P

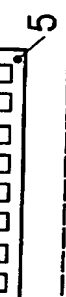
- 24 -

Zusammenfassung:Fußschalter für einen Computer

- 5 Bei einer mit einem Fuß betätigbare Eingabeeinrichtung (3) für einen Computer (1), der eine Diktiereinrichtung bildet, mit Fußeingabemitteln (24) zum manuellen Eingeben einer Steuerinformation (SI), mit der eine Audiowiedergabe-Betriebsart der Diktiereinrichtung aktivierbar (25) und deaktivierbar (26) ist, wobei bei aktivierter Audiowiedergabe-Betriebsart eine in der Diktiereinrichtung gespeicherte Audioinformation (AI) als analoges
- 10 Audiosignal (AS3) an einen Kopfhörer (34) bzw. einen Lautsprecher (32) abgebar ist, und mit Anschlußmitteln (28, 33, 35) zum Anschließen der Eingabeeinrichtung (3) an den Computer (1), wobei von den Anschlußmitteln (28, 33, 35) die Steuerinformation (SI) an den Computer (1) abgebar ist, sind nunmehr die Anschlußmittel (28, 33, 35) zum Empfangen der Audioinformation (AI) als digitale Audiodaten von dem Computer (1) und
- 15 zum Abgeben der Steuerinformation (SI) an den Computer (1) über eine digitale Datenbusverbindung ausgebildet und ist an die Anschlußmittel (28, 33, 35) der Kopfhörer (34) bzw. der Lautsprecher anschließbar.
- (Figur 3)

THIS PAGE BLANK (USPTO)

164



2/3

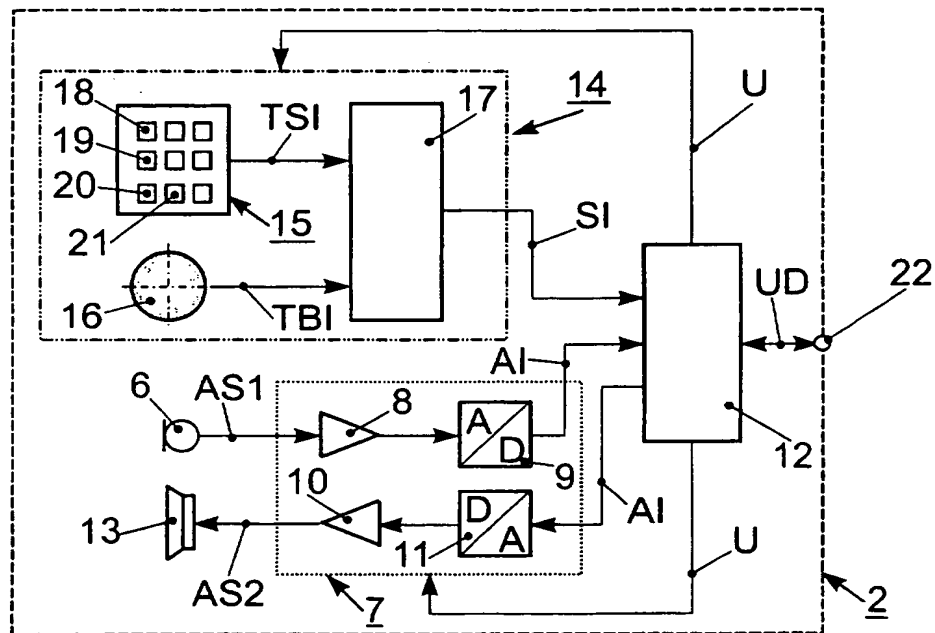


FIG. 2

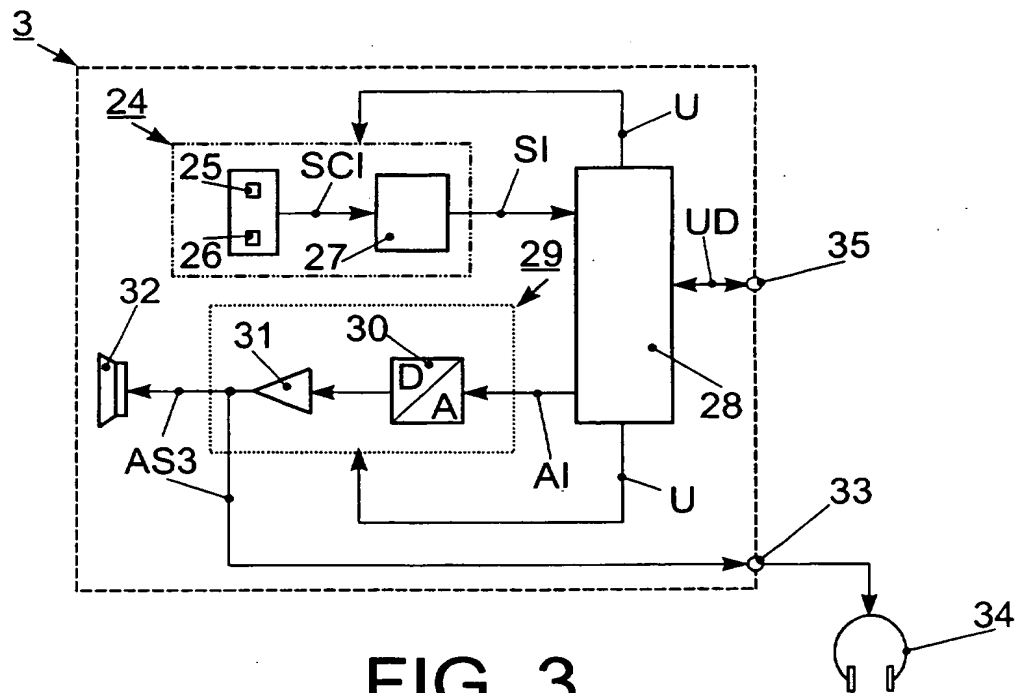


FIG. 3

3/3

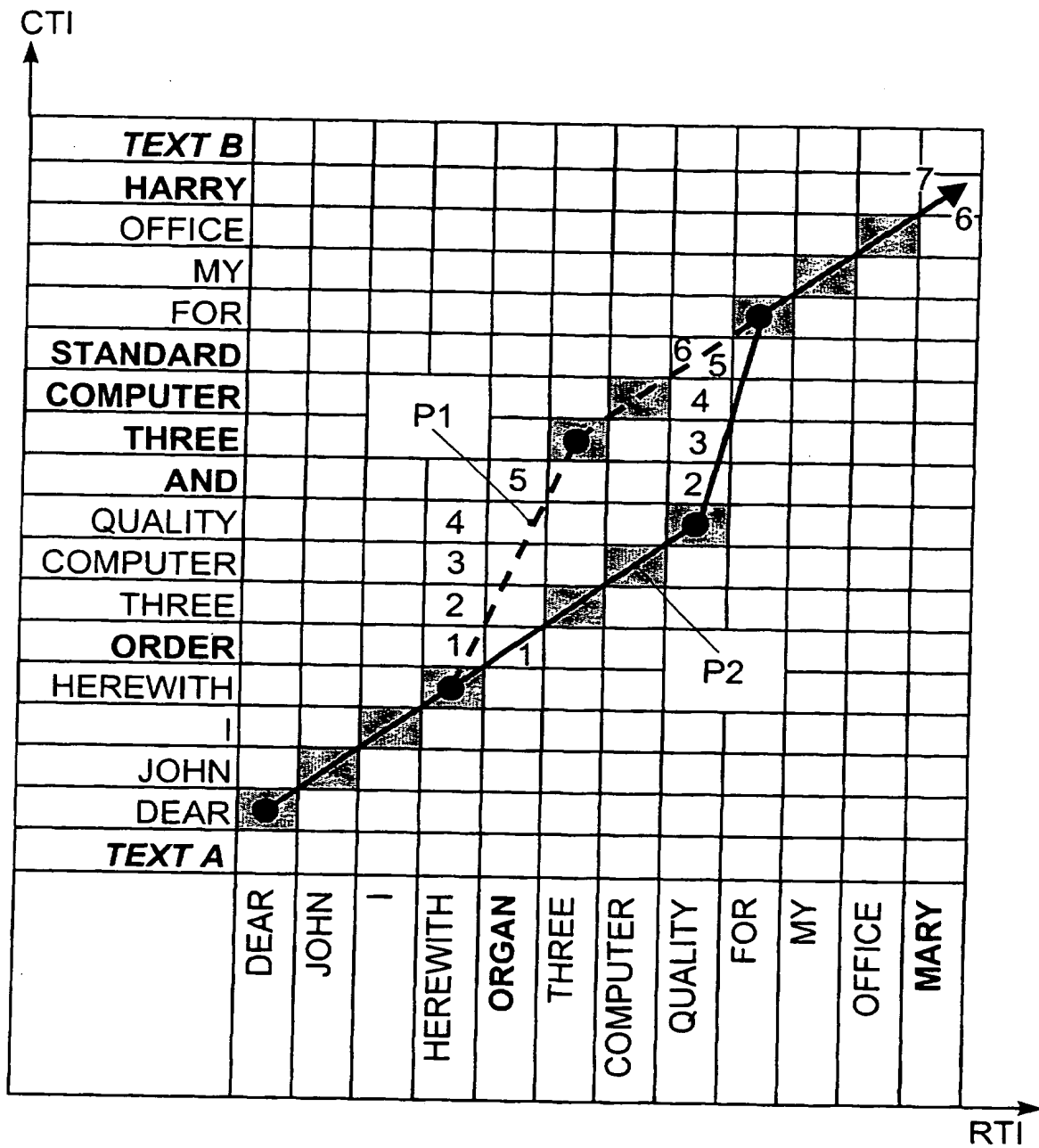


FIG. 4

53

THIS PAGE BLANK (USPTO)
